Multi-Access MIMO Systems with Finite Rate Channel State Feedback

Wei Dai, Brian Rider and Youjian Liu
dai@colorado.edu, brider@euclid.colorado.edu, and eugeneliu@ieee.org
University of Colorado at Boulder

Abstract

This paper characterizes the effect of finite rate channel state feedback on the sum rate of a multi-access multiple-input multiple-output (MIMO) system. We propose to control the users jointly, specifically, we first choose the users jointly and then select the corresponding beamforming vectors jointly. To quantify the sum rate, this paper introduces the *composite Grassmann manifold* and the *composite Grassmann matrix*. By characterizing the distortion rate function on the composite Grassmann manifold and calculating the logdet function of a random composite Grassmann matrix, a good sum rate approximation is derived. According to the distortion rate function on the composite Grassmann manifold, the loss due to finite beamforming decreases exponentially as the feedback bits on beamforming increases.

Index Terms

multi-access, MIMO, limited feedback

I. Introduction

This paper considers the uplink of a cellular system with one base station and multiple users, where both the base station and each user are equipped with multiple antennas. Multiple antenna systems, also known as multiple-input multiple-output (MIMO) systems, provide significant benefit over single antenna systems in terms of either higher spectral efficiency or better reliability. For the uplink of a cellular system, it is reasonable to assume that the base station has the full knowledge about the uplink channel while the users has partial information about the uplink channel through a feedback link from the base station. In practice, it is also reasonable to assume that the feedback link is rate limited.

The purpose of this paper is to quantify the effect of the finite rate channel state feedback on the sum rate. The effect of finite rate feedback on single user MIMO systems are well studied. MIMO systems with only one on-beam are considered in [1] and [2] while systems with multiple on-beams are discussed in [3]–[8]. In the recent works [7] and [8], the effect of finite rate feedback is accurately quantified by characterizing the distortion rate function in the Grassmann manifold. For multi-access systems, the throughput capacity region is characterized in [9] with the assumption that each user has only one antenna and the full channel information is available to all users.

To characterize the feedback gain, we propose to control the users jointly. An simple extension of [10] can show that the optimal strategy is to select the covariance matrices of the transmit signals of the users jointly. It is different from the current systems where the base station controls the users individually. The gain of joint control over individual control is analogous to that of vector quantization over scalar quantization. However, it is difficult to either implement or analyze the fully joint control. For simplicity, this paper proposes a suboptimal strategy employing power on/off strategy, where we first choose the on-users jointly and then select the beamforming vectors jointly. The effect of user choice can be analyzed by extreme order statistics. To quantify the effect of beamforming, the *composite Grassmann manifold* is introduced in this paper. By characterizing the distortion rate function on the composite Grassmann manifold and calculating the logdet function of a random *composite Grassmann matrix*, a good sum rate approximation is derived. According to the distortion rate function on the composite Grassmann manifold, the loss of finite beamforming decreases exponentially as the feedback bits on beamforming increases.

II. System Model

Assume that there are L_R antennas at the base station and N users communicating with the base station. Assume that the user i has $L_{T,i}$ antennas $1 \leq i \leq N$. In this paper, we let $L_{T,i} = L_{T,j} = L_T$ for $1 \leq i, j \leq N$. The signal transmission model is

$$\mathbf{Y} = \sum_{i=1}^{N} \mathbf{H}_i \mathbf{T}_i + \mathbf{W},$$

where $\mathbf{Y} \in \mathbb{C}^{L_R \times 1}$ is the received signal at the base station, $\mathbf{H}_i \in \mathbb{C}^{L_R \times L_T}$ is the channel state matrix for user i, $\mathbf{T}_i$ is the transmitted Gaussian signal vector for user i and $\mathbf{W} \in \mathbb{C}^{L_R \times 1}$ is the additive Gaussian noise vector with zero mean and covariance matrix $\mathbf{I}_{L_R}$. In this paper, we assume the Rayleigh fading channel model, i.e., the entries of $\mathbf{H}_i$ are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian variables with zero mean and unit variance ($\mathcal{CN}(0,1)$) and $\mathbf{H}_i$'s are independent for each channel use.

We assume that there exists a common feedback link from the base station to all the users. At the beginning of each channel use, the channel states $\mathbf{H}_i$'s are perfectly estimated at the receiver. A message, which is a function of the channel state, is sent back to all users through a feedback channel. The feedback is error-free and rate limited. The feedback directs the users to choose their Gaussian signal covariance matrices. In multi-access system, users are uncoordinated. It is reasonable to assume that $\mathrm{E}\left[\mathbf{T}_i \mathbf{T}_j^\dagger\right] = \mathbf{0}$. Let $\mathbf{T} = \left[\mathbf{T}_1^\dagger \cdots \mathbf{T}_N^\dagger\right]^\dagger$ be the overall transmitted Gaussian signal for all users and $\mathbf{\Sigma} \triangleq \mathrm{E}\left[\mathbf{T}\mathbf{T}^\dagger\right]$ be the overall signal covariance matrix. Then $\mathbf{\Sigma}$ is an $NL_T \times NL_T$ block diagonal matrix whose i^{th} diagonal block is the $L_T \times L_T$ covariance matrix $\mathrm{E}\left[\mathbf{T}_i \mathbf{T}_i^\dagger\right]$. Assume there is a covariance matrix codebook $\mathcal{B}_{\mathbf{\Sigma}} = \{\mathbf{\Sigma}_1, \cdots, \mathbf{\Sigma}_{K_\mathcal{B}}\}$ declared to both the base station and the users, where each $\mathbf{\Sigma}_k$ is a proper overall signal covariance matrix and $K_\mathcal{B}$ is the size of the codebook. Let $\mathbf{H} = [\mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_N]$ be the overall channel state matrix. The feedback function φ is a mapping from $\left\{\mathbf{H} \in \mathbb{L}^{L_R \times NL_T}\right\}$ into the index set $\{1, \cdots, K_\mathcal{B}\}$. Subjected to the finite rate feedback constraint

$$|\mathcal{B}_{\mathbf{\Sigma}}| \leq K_\mathcal{B}$$

and the average transmission power constraint

$$\mathrm{E}_{\mathbf{H}}\left[\mathrm{tr}\left(\mathbf{\Sigma}_{\varphi(\mathbf{H})}\right)\right] \leq \rho,$$

we are interested in characterizing the sum rate

$$\max_{\mathcal{B}_{\mathbf{\Sigma}}} \max_{\varphi(\cdot)} \mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_{L_R} + \mathbf{H}\mathbf{\Sigma}_{\varphi(\mathbf{H})}\mathbf{H}^\dagger\right|\right]. \tag{1}$$

Since the variance of the Gaussian noise is normalized, the average power constraint ρ is also the average received signal-to-noise ratio (SNR).

III. Mathematical Preliminary Results

For compositional clarity, this section assembles the useful mathematical results that we derive for later analysis. Due to the space limit, *we omit all the proofs*.

A. Extreme Order Statistics for Chi-Square Random Variable

Let $X_i = \sum_{j=1}^{L} |h_{i,j}|^2$ where $h_{i,j}$ $1 \leq j \leq L$, $1 \leq i \leq n$ are i.i.d. circularly symmetric complex Gaussian variables with zero mean and unit variance. Let us rearrange these i.i.d. chi-square random variables $X_1, \cdots, X_n$ into a nondecreasing sequence $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_n}$. Let n approach infinity, the following theorem gives a formula for $\mathrm{E}\left[\sum_{k=1}^{l} X_{i_{n-k+1}}\right]$ where l is a fixed positive integer.

Theorem 1: Let $X = \sum_{j=1}^{L} |h_j|^2$ where $h_j \sim \mathcal{CN}(0,1)$. Denote the distribution function of X by $F_X(x)$. Then for any fixed positive integer l,

$$\lim_{n\to+\infty} \mathrm{E}\left[\frac{\sum_{k=1}^{l} X_{i_{n-k+1}} - l a_n}{b_n}\right] = l\left(\mu_1^x + 1 - \sum_{k=1}^{l}\frac{1}{k}\right),$$

where a_n is the solution of

$$a_n = \inf\left\{x:\ 1 - F_X(x) \le \frac{1}{n}\right\},$$

$$b_n = \frac{\sum_{i=0}^{L-1}\frac{L-i}{i!}a_n^i}{\sum_{i=0}^{L-1}\frac{1}{i!}a_n^i},$$

and

$$\mu_1^x = \int_{-\infty}^{+\infty} x d e^{-e^{-x}} = 0.577216\cdots.$$

Although this theorem is for asymptotically large n, it gives an accurate approximation when $0 < l \ll n$.

B. Conditioned Eigenvalues of the Wishart Matrix

Let $\mathbf{H} \in \mathbb{L}^{n\times m}$ be a random $n \times m$ matrix whose entries are i.i.d. Gaussian random variables with zero mean and unit variance, where $\mathbb{L}$ is either $\mathbb{R}$ or $\mathbb{C}$ and $m \le n$ w.l.o.g.. The random matrix $\mathbf{W} = \mathbf{H}^\dagger\mathbf{H}$ is Wishart distributed and its distribution is denoted by $W_m(n, \mathbf{I}_m)$.

For a $\mathbf{W} \sim W_m(n, \mathbf{I}_m)$, the following proposition shows that conditioned on the trace, the conditional expectation of a specific eigenvalue of $\mathbf{W}$ is proportional to the condition with a ratio independent of that condition.

Proposition 1: Let $\mathbf{W} \sim W_m(n, \mathbf{I}_m)$ where $n \ge m$. List the ordered eigenvalues of $\mathbf{W}$ as $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m \ge 0$. Then conditioned on the trace of $\mathbf{W}$, i.e., $\sum_{i=1}^{m}\lambda_i = c$ where $c > 0$, the ratio between the conditional expectation of λ_i and the condition c is a constant ζ_i independent of c, i.e.,

$$\mathrm{E}\left[\lambda_i \middle| \sum_{i=1}^{m}\lambda_i = c\right] = \zeta_i c$$

where

$$\zeta_i = \frac{\int_{\substack{\sum\lambda_j=1\\ \lambda_1\ge\cdots\ge\lambda_m}} \lambda_i \prod_{j=1}^{m}\lambda_i^{\frac{\beta}{2}(n-m+1)-1}\left|\Delta_m(\lambda)\right|^\beta \prod_{j=1}^{m} d\lambda_j}{\int_{\substack{\sum\lambda_j=1\\ \lambda_1\ge\cdots\ge\lambda_m}} \prod_{j=1}^{m}\lambda_i^{\frac{\beta}{2}(n-m+1)-1}\left|\Delta_m(\lambda)\right|^\beta \prod_{j=1}^{m} d\lambda_j},$$

$\beta = 1$ if $\mathbb{L} = \mathbb{R}$ or $\beta = 2$ if $\mathbb{L} = \mathbb{C}$, and $|\Delta_m(\lambda)| = \prod_{i<j}^{m}(\lambda_i - \lambda_j)$.

In general, it is not easy to calculate the constant ζ_i $1 \le i \le m$. Fortunately, the constants can be well approximated by asymptotics. Due to the space limit, we only present the asymptotic formula for ζ_1 in the following proposition.

Proposition 2: Let the random matrix $\mathbf{W} \sim W_m(n, \mathbf{I}_m)$ where $n \ge m$. Define $y \triangleq \frac{m}{n}$. Then the asymptotic approximation gives

$$\mathrm{E}\left[\lambda_1 \middle| \sum_{i=1}^{m}\lambda_i = c\right] \approx \frac{1}{\pi}\left[\pi - a + \frac{1}{2}\sin(2a)\right]c,$$

where a satisfies

$$\frac{1}{m} = \begin{cases} \frac{1}{\pi}\left[\pi - a - \frac{1}{\sqrt{y}}\sin(a) + \frac{1-y}{y}\theta_y\right] & \text{if } y < 1 \\ \frac{1}{\pi}\left[\pi - a - \sin(a)\right] & \text{if } y = 1 \end{cases},$$

and

$$\theta_y = \tan^{-1}\left(\frac{\sqrt{y}\sin(a)}{1 - \sqrt{y}\cos(a)}\right).$$

C. The Grassmann Manifold and the Composite Grassmann Manifold

The Grassmann manifold is the geometric object relevant to the beamforming quantization analysis. The Grassmann manifold $\mathcal{G}_{n,m}(\mathbb{L})$ is the set of m-dimensional planes (passing through the origin) in Euclidean n-space $\mathbb{L}^n$. A generator matrix $\mathbf{P} \in \mathbb{L}^{n\times m}$ for an m-plane $P \in \mathcal{G}_{n,m}(\mathbb{L})$ is the matrix whose columns are orthonormal and span P. The generator matrix is not unique. That is, if $\mathbf{P}$ generates P then $\mathbf{PU}$ also generates P for any $m \times m$ orthogonal/unitary matrix $\mathbf{U}$ (w.r.t. $\mathbb{L} = \mathbb{R}/\mathbb{C}$ respectively) [11]. The chordal distance between two m-planes $P_1, P_2 \in \mathcal{G}_{n,m}(\mathbb{L})$ can be defined by their generator matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ via $d_c(P_1, P_2) = \frac{1}{\sqrt{2}}\left\|\mathbf{P}_1\mathbf{P}_1^\dagger - \mathbf{P}_2\mathbf{P}_2^\dagger\right\|_F$ [11]. The uniform distribution on $\mathcal{G}_{n,m}(\mathbb{L})$ with density function $f_P(\cdot)$ satisfies $f_P(P_1) = f_P(P_2)$ for arbitrary $P_1, P_2 \in \mathcal{G}_{n,m}(\mathbb{L})$ [12].

For quantizations on $\mathcal{G}_{n,m}(\mathbb{L})$, the corresponding distortion rate function has been characterized [7]. A quantization q on $\mathcal{G}_{n,m}(\mathbb{L})$ is a mapping from $\mathcal{G}_{n,p}(\mathbb{L})$ to a subset of $\mathcal{G}_{n,p}(\mathbb{L})$, which is typically called a code $\mathcal{C}$, i.e., $q : \mathcal{G}_{n,p}(\mathbb{L}) \to \mathcal{C}$. Define the distortion metric as the squared chordal distance. Then the distortion associated with a quantization q is

$$D \triangleq \mathrm{E}_Q\left[d_c^2(Q, q(Q))\right],$$

where the source Q is randomly distributed in $\mathcal{G}_{n,m}(\mathbb{L})$. Assume that the source Q is uniformly distributed in $\mathcal{G}_{n,p}(\mathbb{L})$. For any given code $\mathcal{C}$, the optimal quantization to minimize the distortion is[1]

$$q(Q) = \arg\min_{P\in\mathcal{C}} d_c(P, Q).$$

The distortion associated with this quantization is

$$D(\mathcal{C}) = \mathrm{E}_Q\left[\min_{P\in\mathcal{C}} d_c^2(P, Q)\right].$$

For a given code size K where K is a positive integer, the distortion rate function is[2]

$$D^*(K) = \inf_{\mathcal{C}:|\mathcal{C}|=K} D(\mathcal{C}).$$

In [8], we derive a lower bound and an upper bound for $\mathbb{L} = \mathbb{C}$

$$\frac{t}{t+1}\eta^{-\frac{1}{t}}2^{-\frac{\log_2 K}{t}} \lesssim D^*(K) \lesssim \frac{\Gamma\left(\frac{1}{t}\right)}{t}\eta^{-\frac{1}{t}}2^{-\frac{\log_2 K}{t}},$$

where $t = m(n-m)$,

$$\eta = \begin{cases} \frac{1}{t!}\prod_{i=1}^{m}\frac{(n-i)!}{(m-i)!} & \text{if } 1 \le m \le \frac{n}{2} \\ \frac{1}{t!}\prod_{i=1}^{n-m}\frac{(n-i)!}{(n-m-i)!} & \text{if } \frac{n}{2} \le m \le n \end{cases},$$

and the symbol $\lesssim$ denotes the *main order inequality*, $f(K) \lesssim g(K)$ if $\lim_{K\to+\infty}\frac{f(K)}{g(K)} \le 1$.

To treat multi-access MIMO systems, we define the *composite Grassmann manifold*. The k-composite Grassmann manifold $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$ is a Cartesian product of k $\mathcal{G}_{n,m}(\mathbb{L})$'s. Denote $P^{(k)}$ an element in $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$.

$$P^{(k)} = (P_1, \cdots, P_k)$$

where $P_i \in \mathcal{G}_{n,m}(\mathbb{L})$ $1 \le i \le k$. For any $P_1^{(k)}, P_2^{(k)} \in \mathcal{G}_{n,m}^{(k)}(\mathbb{L})$, we define the chordal distance between them

$$d_c\left(P_1^{(k)}, P_2^{(k)}\right) = \sqrt{\sum_{i=1}^{k} d_c^2(P_{1,i}, P_{2,i})},$$

[1]The ties, i.e., the case that $\exists P_1, P_2 \in \mathcal{C}$ such that $d_c(P_1, Q) = \min_{P\in\mathcal{C}} d_c(P, Q) = d_c(P_2, Q)$, are broken arbitrarily because the probability of ties is zero.

[2]The standard definition of the distortion rate function is a function of the code rate defined by $\log_2 K$. The definition in this paper is equivalent to the standard one.

where $P_1^{(k)} = (P_{1,1}, \cdots, P_{1,k})$ and $P_2^{(k)} = (P_{2,1}, \cdots, P_{2,k})$. It is easy to verify that the chordal distance on $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$ is well defined.

This paper characterizes the distortion rate function for quantizations on $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$. Define the distortion metric on $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$ as the square chordal distance on it. Assume a uniformly distributed source $Q^{(k)}$ in $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$. The following theorem characterizes the distortion rate function for quantizations on $\mathcal{G}_{n,m}^{(k)}(\mathbb{C})$.

Theorem 2: The distortion rate function on $\mathcal{G}_{n,m}^{(k)}(\mathbb{C})$ is upper bounded and lower bounded by

$$\frac{kt}{kt+1}\left(\frac{\Gamma^k(t+1)}{\Gamma(kt+1)}\eta^k\right)^{-\frac{1}{kt}} 2^{-\frac{\log_2 K}{kt}} \lesssim D^*(K) \lesssim \frac{\Gamma\left(\frac{1}{kt}\right)}{kt}\left(\frac{\Gamma^k(t+1)}{\Gamma(kt+1)}\eta^k\right)^{-\frac{1}{kt}} 2^{-\frac{\log_2 K}{kt}},$$

where $t = m(n-m)$,

$$\eta = \begin{cases} \frac{1}{t!}\prod_{i=1}^{m}\frac{(n-i)!}{(m-i)!} & \text{if } 1 \le m \le \frac{n}{2} \\ \frac{1}{t!}\prod_{i=1}^{n-m}\frac{(n-i)!}{(n-m-i)!} & \text{if } \frac{n}{2} \le m \le n \end{cases},$$

and the symbol $\lesssim$ denotes the *main order inequality*, $f(K) \lesssim g(K)$ if $\lim\limits_{K\to+\infty}\frac{f(K)}{g(K)} \le 1$.

It is noteworthy that the upper bound is derived by computing the average distortion over the ensemble of random codes. In practice, we often use the upper bound as an approximation to the actual distortion rate function.

D. Composite Grassmann Matrix

Roughly speaking, a *composite Grassmann matrix* is the generator matrix for an element in $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$. Let $P^{(k)} = (P_1, \cdots, P_k) \in \mathcal{G}_{n,m}^{(k)}(\mathbb{L})$. The composite matrix $\mathbf{P}^{(k)}$ generating $P^{(k)}$ is $\mathbf{P}^{(k)} = [\mathbf{P}_1 \cdots \mathbf{P}_k]$ where $\mathbf{P}_1, \cdots, \mathbf{P}_k$ are the generator matrices for $P_1, \cdots, P_k$ respectively. Since the generator matrix for a plane in the Grassmann manifold is not unique, the composite Grassmann matrix generating $P^{(k)}$ is not unique either. Let $\mathbf{P}^{(k)}$ be a generator matrix for $P^{(k)}$. The matrix $\mathbf{P}^{(k)}\mathbf{U}^{(k)}$, where $\mathbf{U}^{(k)}$ is the arbitrary $km \times km$ block diagonal matrix whose k diagonal blocks are $m \times m$ orthogonal/unitary matrices (w.r.t. $\mathbb{L} = \mathbb{R}/\mathbb{C}$ respectively), also generates $P^{(k)}$. In this paper, the set of composite Grassmann matrices for $\mathcal{G}_{n,m}^{(k)}(\mathbb{L})$ is denoted by $\mathcal{M}_{n,m}^{(k)}(\mathbb{L})$.

For a random composite Grassmann matrix $\mathbf{P}^{(k)}$, the following theorem bounds $\mathrm{E}\left[\log\left|\mathbf{I} + c\mathbf{P}^{(k)\dagger}\mathbf{P}^{(k)}\right|\right]$.

Theorem 3: Let $\mathbf{P}^{(k)} \in \mathcal{M}_{n,1}^{(k)}(\mathbb{L})$ be uniformly distributed. For any positive constant c,

$$\begin{aligned} \mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_k + \frac{c}{n}\mathbf{H}^\dagger\mathbf{H}\right|\right] &\le \mathrm{E}_{\mathbf{P}^k}\left[\log\left|\mathbf{I}_k + c\mathbf{P}^{(k)\dagger}\mathbf{P}^{(k)}\right|\right] \\ &\le \log \mathrm{E}_{\mathbf{P}^k}\left[\left|\mathbf{I}_k + c\mathbf{P}^{(k)\dagger}\mathbf{P}^{(k)}\right|\right], \end{aligned}$$

where $\mathbf{H} \in \mathbb{L}^{n\times k}$ has i.i.d. Gaussian entries with zero mean and unit variance.

In the above theorem, both bounds can be computed explicitly. In [13], we derive an asymptotic formula to approximate the lower bound. Let n and k approach infinity simultaneously with fixed ratio,

$$\begin{aligned} &\lim_{(n,k)\to+\infty}\frac{1}{\min(n,k)}\mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_k + \frac{c}{n}\mathbf{H}^\dagger\mathbf{H}\right|\right] \\ &= \log(w) - \log(\alpha) - \frac{u}{r} - \frac{(1-y)\log(1-ur)}{y}, \end{aligned}$$

where $y \triangleq \frac{\min(n,k)}{\max(n,k)}$, $r \triangleq \sqrt{y}$, $\alpha \triangleq \frac{n}{\min(n,k)\cdot c}$, $w \triangleq \frac{1}{2}\left(1 + y + \alpha + \sqrt{(1+y+\alpha)^2 - 4y}\right)$ and $u \triangleq \frac{1}{2r}\left(1 + y + \alpha - \sqrt{(1+y+\alpha)^2 - 4y}\right)$. Formulas for the upper bound are also derived in this paper. Due to the space limit, we only present the formulas for $1 \le k \le 5$. The expectation $\mathrm{E}_{\mathbf{P}^k}\left[\left|\mathbf{I}_k + c\mathbf{P}^{(k)\dagger}\mathbf{P}^{(k)}\right|\right]$ can be calculated by

k=1 $1+c$;
k=2 $(1+c)^2 - c^2 \frac{1}{n}$;
k=3 $(1+c)^3 - c^2 (1+c) \frac{3}{n} + c^3 \frac{2}{n^2}$;
k=4 $(1+c)^4 - c^2 (1+c)^2 \frac{6}{n} + c^3 (1+c) \frac{8}{n^2} - c^4 \left(\frac{6}{n^3} - \frac{3}{n^2}\right)$; and
k=5 $(1+c)^5 - c^2 (1+c)^3 \frac{10}{n} + c^3 (1+c)^2 \frac{20}{n^2} - c^4 (1+c) \left(\frac{30}{n^3} - \frac{15}{n^2}\right) + c^5 \left(\frac{24}{n^4} - \frac{20}{n^3}\right)$.

IV. The Suboptimal Strategy and the Sum Rate

This section is devoted to calculate the sum rate of a multi-access MIMO system with finite rate feedback. The computation of the sum rate (1) involves two correlated optimization problems: one is with respect to the feedback function φ and the other optimization is over all possible covariance matrix codebooks. The direct calculation of (1) is difficult.

To reduce the complexity, we propose a suboptimal strategy to control the users jointly. Specifically, we first choose the on-users jointly and then select the corresponding beamforming vectors jointly. It is different from the current system where users are controlled individually.

The assumptions for transmission are as follows.

T1) Power on/off strategy. In power on/off strategy, the user i's covariance matrix is of the form $\mathbf{\Sigma}_i = P_{\text{on}} \mathbf{Q}_i \mathbf{Q}_i^\dagger$, where P_{on} is a fixed positive constant to denote on-power and $\mathbf{Q}_i$ is the beamforming matrix for user i. Denote each column of $\mathbf{Q}_i$ an *on-beam* and the number of the columns of $\mathbf{Q}_i$ by l_i, then $\mathbf{Q}_i^\dagger \mathbf{Q}_i = \mathbf{I}_{l_i}$ where $0 \le l_i \le L_T$ and $l_i = 0$ is for the case that the user i is off. This assumption is motivated by the fact that power on/off strategy is near-optimal for single user MIMO systems [7].

T2) At most one on-beam per user. This assumption implies either $l_i = 0$ or $l_i = 1$. It is proposed so that each user has larger probability to be turned on.

T3) Constant number of on-beams for a given SNR. Let $l = \sum_{i=1}^N l_i$ be the total number of on-beams. we assume that l is a constant independent of the specific channel realization for a given SNR. This assumption is motivated by the fact that constant number of on-beams is near optimal for single user systems [7]. It will be validated for multi-access systems in later analysis.

The feedback is described as below.

F1) User selection criterion. Assume that l users will be turned on. We choose the l users with the largest channel state Frobenius norms, i.e. $\left\|\mathbf{H}_{i_j}\right\| \ge \left\|\mathbf{H}_i\right\|$ for all $i \notin \{i_j : 1 \le j \le l\}$ where $\|\cdot\|$ is the Frobenius norm and $i_1, \cdots, i_l$ are the users chosen to be on (*on-users*).

According to this user selection criterion, the feedback contains two parts, one of which indicates the l on-users and the other of which is for beamforming. Let $i_1, \cdots, i_l$ be the on-users and $\mathbf{b}_1, \cdots, \mathbf{b}_l$ be the beamforming vectors for those users. Then $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_l] \in \mathcal{M}_{L_T,1}^{(l)}(\mathbb{C})$ where $\mathcal{M}_{L_T,1}^{(l)}(\mathbb{C})$ is the set of composite Grassmann matrix (Section III-D). Denote the beamforming codebook $\mathcal{B} = \left\{\mathbf{B}_k : \mathbf{B}_k \in \mathcal{M}_{L_T,1}^{(l)}(\mathbb{C}), 1 \le k \le |\mathcal{B}|\right\}$. Then the overall feedback codebook is the Cartesian product of the set of on-users $\{(i_1, \cdots, i_l)\}$ and the beamforming codebook $\mathcal{B}$. Let $\mathbf{H}_{i_j}$ be the channel state matrix for the user i_j and $\mathbf{v}_{j,1}$ be the right singular vector corresponding to the largest singular value of $\mathbf{H}_{i_j}$. Define $\mathbf{V} \triangleq [\mathbf{v}_{1,1} \cdots \mathbf{v}_{l,1}]$. The beamforming feedback function is defined as the following.

F2) Beamforming feedback function.

$$\begin{aligned} \varphi\left([\mathbf{H}_{i_1} \cdots \mathbf{H}_{i_l}]\right) &\triangleq \underset{1 \le k \le |\mathcal{B}|}{\arg\min}\, d_c^2(\mathbf{V}, \mathbf{B}_k) \\ &= \underset{1 \le k \le |\mathcal{B}|}{\arg\max} \sum_{j=1}^{l} \left|\mathbf{v}_{j,1}^\dagger \mathbf{b}_{k,j}\right|^2, \end{aligned} \tag{2}$$

where $d_c^2(\mathbf{V}, \mathbf{B}_k)$ denotes the chordal distance between the elements in the composite Grassmann manifold $\mathcal{G}_{L_T,1}^{(l)}(\mathbb{C})$ generated by $\mathbf{V}$ and $\mathbf{B}_k$, and $\mathbf{b}_{k,j}$ is the j^{th} column of the k^{th} beamforming matrix $\mathbf{B}_k \in \mathcal{B}$.

The feedback assumptions F1 and F2 will be validated in the later analysis.

The above assumptions define a suboptimal strategy for multi-access MIMO systems. The key point is that the user choice is independent of the channel directions and the beamforming is independent of the channel strengths (norms). In this way, the effect of user choice and beamforming can be studied separately. Before diving into the general analysis, we discuss a special case, antenna selection, to get some intuition.

A. Antenna Selection

The system model for antenna selection is

$$\mathbf{Y} = \sum_{i=1}^{NL_T} \mathbf{h}_i T_i + \mathbf{W},$$

where $\mathbf{h}_i$ is the i^{th} column of the overall channel state matrix $\mathbf{H}$. For each channel realization $\mathbf{H}$, we simply choose l antennas $i_1, \cdots, i_l$ such that $\left\|\mathbf{h}_{i_j}\right\| \geq \|\mathbf{h}_i\|$ for all $i \notin \{i_j : 1 \leq j \leq l\}$. Here, we actually do not require one on-beam per on-user (Assumption T2). Write $\mathbf{h}_{i_j} = n_j \xi_j$ where n_j is the Frobenius norm of $\mathbf{h}_{i_j}$ and ξ_j is the unit vector to present the direction of $\mathbf{h}_{i_j}$. Define $\mathbf{\Xi} \triangleq [\xi_1 \cdots \xi_l]$. We have the following upper bound on the sum rate.

$$\begin{aligned}
&\mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_{L_R} + \frac{\rho}{l}\sum_{j=1}^{l}\mathbf{h}_{i_j}\mathbf{h}_{i_j}^{\dagger}\right|\right] \\
&= \mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_l + \frac{\rho}{l}\mathrm{diag}\left[n_1^2, \cdots, n_l^2\right]\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right] \\
&\leq \mathrm{E}_{\mathbf{\Xi}}\left[\log\left|\mathbf{I}_l + \frac{\rho}{l}\mathrm{E}_{\mathbf{n}^2}\left[\mathrm{diag}\left[n_1^2, \cdots, n_l^2\right]\right]\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right] \\
&= \mathrm{E}_{\mathbf{\Xi}}\left[\log\left|\mathbf{I}_l + \frac{\rho}{l}\frac{\mathrm{E}_{\mathbf{n}^2}\left[\sum_{j=1}^{l} n_j^2\right]}{l}\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right],
\end{aligned} \tag{3}$$

where the inequality follows from the concavity of $\log|\cdot|$ function and the fact that n_j^2's and $\mathbf{\Xi}$ are independent. Noting that $\|\mathbf{h}_i\|^2$'s are i.i.d. chi-square random variables, an accurate approximation to $\mathrm{E}_{\mathbf{n}^2}\left[\sum_{j=1}^{l} n_j^2\right]$ can be obtained for $l \ll N$ by applying the asymptotic extreme order statistics in Theorem 1. On the other hand, it can be proved that ξ_j's are independent and uniformly distributed unit vectors. Regarding $\mathbf{\Xi}$ as a Cartesian product of ξ_j's, $\mathbf{\Xi}$ is also uniformly distributed in $\mathcal{M}_{L_R,1}^{(l)}$, the set of composite Grassmann matrix. According to the results in Section III-D for $\mathrm{E}\left[\log\left|\mathbf{I} + c\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right]$, the upper bound of the sum rate (3) can be characterized. Simulation show that the upper bound (3) is tight. The sum rate of antenna selection is then approximately characterized.

B. General Beamforming

With the assumptions T1-3, F1 and F2, the signal model for the general beamforming is

$$\mathbf{Y} = \sum_{j=1}^{l} \mathbf{H}_{i_j}\mathbf{b}_{\varphi(\mathbf{H}),j}T_j + \mathbf{W},$$

where $\mathbf{b}_{\varphi(\mathbf{H}),j}$ is the j^{th} column of the feedback beamforming matrix $\mathbf{B}_{\varphi(\mathbf{H})} \in \mathcal{B}$. For notational convenience, we denote $\mathbf{b}_{\varphi(\mathbf{H}),j}$ by $\mathbf{b}_j^*$ and the equivalent channel vector $\mathbf{H}_{i_j}\mathbf{b}_j^*$ by $\hat{\mathbf{h}}_j$. Let n_j be the Frobenius

norm of $\hat{\mathbf{h}}_j$, ξ_j be the unit vector presenting the direction of $\hat{\mathbf{h}}_j$ and $\mathbf{\Xi} = [\xi_1 \cdots \xi_l]$. Then the sum rate is given by

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_{L_R} + \frac{\rho}{l}\sum_{j=1}^{l}\hat{\mathbf{h}}_j\hat{\mathbf{h}}_j^\dagger\right|\right] \\ &= \mathrm{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_l + \frac{\rho}{l}\mathrm{diag}\left[n_1,\cdots,n_l^2\right]\mathbf{\Xi}^\dagger\mathbf{\Xi}\right|\right]. \end{aligned}$$

It can be proved that ξ_j's are uniformly distributed and independent of n_j's. Denote the singular value decomposition of $\mathbf{H}_{i_j}$ by $\mathbf{U}_j\mathbf{\Lambda}_j\mathbf{V}_j^\dagger$. After beamforming, the equivalent channel vector for user i_j is $\hat{\mathbf{h}}_j = \mathbf{U}_j\left(\mathbf{\Lambda}_j\mathbf{V}_j^\dagger\mathbf{b}_j^*\right) = \mathbf{U}_j\tilde{\xi}_j n_j$, where $\tilde{\xi}_j$ is the direction of the vector $\mathbf{\Lambda}_j\mathbf{V}_j^\dagger\mathbf{b}_j^*$. Since the user choice is only dependent on $\mathbf{\Lambda}_j$'s and the beamforming matrix selection is only relevant to $\mathbf{V}_j$'s, $\mathbf{U}_j$'s are independent and uniformly distributed. According to [14, Thm. 6.1], $\xi_j = \mathbf{U}_j\tilde{\xi}_j$ is uniformly distributed and independent of n_j's. Thus, similar to (3), the sum rate of general beamforming can be upper bounded by

$$\mathrm{E}_{\mathbf{\Xi}}\left[\log\left|\mathbf{I}_l + \frac{\rho}{l}\frac{\mathrm{E}_{\mathbf{n}^2}\left[\sum_{j=1}^{l}n_j^2\right]}{l}\mathbf{\Xi}^\dagger\mathbf{\Xi}\right|\right], \tag{4}$$

where $\mathbf{\Xi} = [\xi_1 \cdots \xi_l]$.

It is more involved to calculate $\mathrm{E}\left[\sum_{j=1}^{l}n_j^2\right]$. Let $\lambda_{j,k}$ $1 \leq k \leq L_T$ be the ordered eigenvalues of $\mathbf{H}_{i_j}^\dagger\mathbf{H}_{i_j}$ such that $\lambda_{j,1} \geq \lambda_{j,2} \geq \cdots \geq \lambda_{j,L_T} \geq 0$. Let $\mathbf{v}_{j,k}$ be the right singular vector of $\mathbf{H}_{i_j}$ corresponding to the k^{th} largest singular value $\sqrt{\lambda_{j,k}}$. Then

$$\begin{aligned} \mathrm{E}\left[n_j^2\right] &= \mathrm{E}\left[\left\|\mathbf{H}_{i_j}\mathbf{b}_j^*\right\|^2\right] = \mathrm{E}\left[\mathbf{b}_j^{*\dagger}\mathbf{H}_{i_j}^\dagger\mathbf{H}_{i_j}\mathbf{b}_j^*\right] \\ &= \mathrm{E}\left[\sum_{k=1}^{L_T}\lambda_{j,k}\left|\mathbf{v}_{j,k}^\dagger\mathbf{b}_j^*\right|^2\right] \\ &= \sum_{k=1}^{L_T}\mathrm{E}\left[\lambda_{j,k}\right]\mathrm{E}\left[\left|\mathbf{v}_{j,k}^\dagger\mathbf{b}_j^*\right|^2\right]. \end{aligned} \tag{5}$$

where the last equality follows from the fact that the beamforming is independent of the channel norms, i.e., $\sum_{k=1}^{L_T}\lambda_{j,k}$'s. The $\mathrm{E}\left[\lambda_{j,k}\right]$'s can be calculated by

$$\mathrm{E}\left[\lambda_{j,k}\right] = \mathrm{E}\left[\mathrm{E}\left[\lambda_{j,k}\left|\left\|\mathbf{H}_{i_j}\right\|^2\right.\right]\right] = \zeta_k\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right], \tag{6}$$

where the last equality is a direct application of Proposition 1 in Section III-B. To evaluate $\mathrm{E}\left[\left|\mathbf{v}_{j,k}^\dagger\mathbf{b}_j^*\right|^2\right]$, we need the following proposition.

Proposition 3: Consider the beamforming feedback function in (2). Define $\gamma \triangleq \mathrm{E}\left[\sum_{j=1}^{l}\left|\mathbf{v}_{j,1}^\dagger\mathbf{b}_j^*\right|^2\right]$. Then $\mathrm{E}\left[\left|\mathbf{v}_{j,1}^\dagger\mathbf{b}_j^*\right|^2\right] = \frac{\gamma}{l}$ and $\mathrm{E}\left[\left|\mathbf{v}_{j,k}^\dagger\mathbf{b}_j^*\right|^2\right] = \left(1-\frac{\gamma}{l}\right)/\left(L_T-1\right)$ for all $1 \leq j \leq l$ and $2 \leq k \leq L_T$.

Apply this proposition and substitute (6) into (5). After some elementary manipulations, we have

$$\mathrm{E}\left[\sum_{j=1}^{l}n_j^2\right] = \left(\frac{\zeta_1\gamma}{l} + \frac{\left(1-\zeta_1\right)\left(l-\gamma\right)}{l\left(L_T-1\right)}\right)\sum_{j=1}^{l}\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right]. \tag{7}$$

Theorem 1 and Proposition 2 provide asymptotic formulas to approximate $\sum_{j=1}^{l}\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right]$ and ζ_1 respectively. Define $K \triangleq |\mathcal{B}|$ the size of the beamforming codebook. The maximum γ achievable $\gamma_{\sup}$

is a function of K. According to the distortion rate function $D^*(K)$ for quantizations on the composite Grassmann manifold $\mathcal{G}_{L_T,1}^{(l)}(\mathbb{C})$,

$$\gamma_{\text{sup}} \triangleq \sup_{\mathcal{B}:\,|\mathcal{B}|\leq K} \gamma = l - D^*(K).$$

Substitute γ_{sup} into (7). The expectation $\mathrm{E}\left[\sum_{j=1}^{l} n_j^2\right]$ can be calculated as a function of K.

Finally, substituting the value of $\mathrm{E}\left[\sum_{j=1}^{l} n_j^2\right]$ into (4) and employing the results in Section III-D for $\mathrm{E}\left[\log\left|\mathbf{I}+c\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right]$, the upper bound of the sum rate (4) can be characterized. Simulations show that this upper bound is tight. The sum rate is therefore approximately characterized.

C. The Effect of Finite Rate Feedback

The above analysis characterizes the effect of finite rate channel state feedback. The upper bound (4) shows that the effect of feedback is quantified by $\mathrm{E}\left[\sum_{j=1}^{l} n_j^2\right]$. Formula (5) shows that the effect of user choice and beamforming can be analyzed separately.

According to (7), the effects of user choice is reflected by $\sum_{j=1}^{l}\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right]$. Maximization of the sum rate requires to maximize $\sum_{j=1}^{l}\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right]$ and thus the user selection criterion (Assumption F1) is validated. Furthermore, the term $\sum_{j=1}^{l}\mathrm{E}\left[\left\|\mathbf{H}_{i_j}\right\|^2\right]$ is an increasing function of the number of users N (Refer to Section III-A). The more users the system has, the larger the sum rate is.

The effect of beamforming can be analyzed according to (7). Define $K \triangleq |\mathcal{B}|$ and $R_{\text{fb}} \triangleq \log_2 K$. Assume that R_{fb} is large so that $(l-\gamma) \ll \gamma$. Then approximately, $\mathrm{E}\left[\sum_{j=1}^{l} n_j^2\right]$ is proportional to γ. The beamforming feedback function should maximize γ and Assumption F2 is therefore verified. Denote $l-\gamma_{\text{sup}} = D^*(K)$ the beamforming loss. From the distortion rate function on the $\mathcal{G}_{L_T,1}^{(l)}(\mathbb{C})$, $l-\gamma_{\text{sup}}$ is a exponentially decreasing function of $R_{\text{fb}}/l\,(L_T-1)$. We expect that a few feedback bits on beamforming could have large gain while more feedback bits wouldn't gain much further.

The assumption T3 about the constant number of on-beams can be validated as well. Assume that both the number of users N and the feedback bits on beamforming R_{fb} are large. Because of the user choice and beamforming, the quantities n_j^2 $1 \leq j \leq l$ are relatively "stable", i.e., the fluctuations of n_j^2's are relatively small. It is reasonable to assume constant number of on-beams for multi-access system.

The antenna selection can be viewed as a special case of general beamforming where the beamforming vector is always a column of the identity matrix. For general beamforming, $\log_2\begin{pmatrix}N\\l\end{pmatrix}+R_{\text{fb}}$ feedback bits are needed. For antenna selection, there are $\log_2\begin{pmatrix}NL_T\\l\end{pmatrix} \approx \log_2\begin{pmatrix}N\\l\end{pmatrix}+l\log_2 L_T$ feedback bits needed. Since antenna selection does not assume one on-beam per on-user (Assumption T2), it is expected that the sum rate of antenna selection is close to but better than that of general beamforming with the same l and $R_{\text{fb}} = l\log_2 L_T$. The improvement is due to the extra freedom the antenna selection has.

D. Simulation

The sum rates of antenna selection and general beamforming are given in Fig. 1 and Fig. 2 respectively. Simulations show that the upper bound (4) (solid lines) is tight. Note that the upper bound (4) is of the form $\mathrm{E}\left[\log\left|\mathbf{I}+c\mathbf{\Xi}^{\dagger}\mathbf{\Xi}\right|\right]$. Theoretical analysis (Theorem 3) gives an upper bound (plus markers) and a lower bound ('x' markers) on (4). Simulations show that these theoretical approximations are accurate.

Fig 2 also depicts the gain of beamforming. The sum rate by finite rate beamforming feedback (circles) is compared to that of perfect beamforming (dash-dot lines). Simulation shows that with several feedback bits on beamforming, the corresponding sum rate is close to that of perfect beamforming. As a special case of general beamforming, antenna selection is shown to be similar to but better than general beamforming with the same l and $R_{\text{fb}} = l\log_2 L_T$.

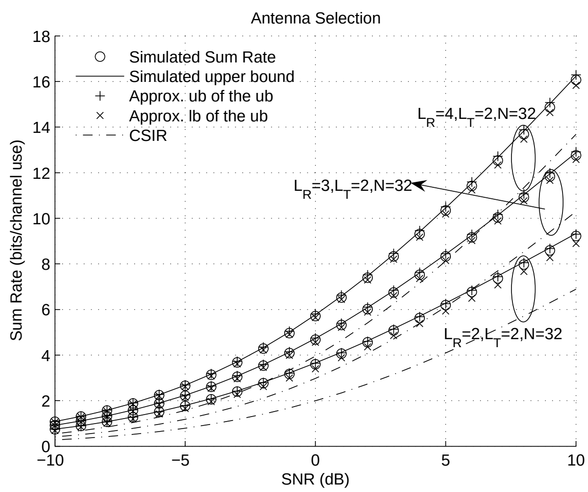

Fig. 1. Sum Rate for Antenna Selection.

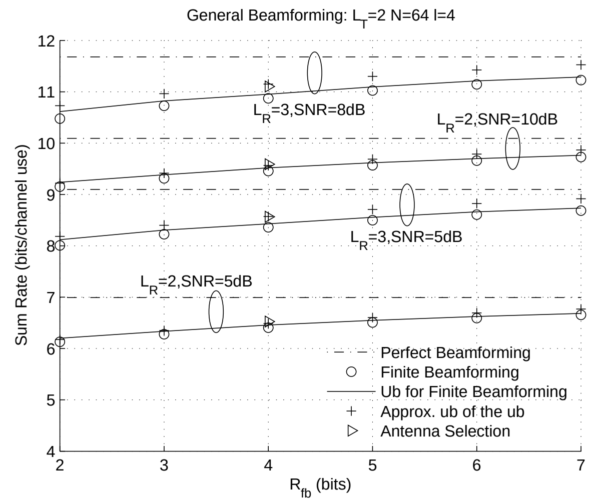

Fig. 2. Sum Rate for General Beamforming.

V. Conclusion

This paper proposes a strategy where users are controlled jointly. The effect of user choice is analyzed by extreme order statistics and the effect of beamforming is quantified by the distortion rate function in the composite Grassmann manifold. By characterizing the distortion rate function on the composite Grassmann manifold and calculating the logdet function of a random composite Grassmann matrix, a good sum rate approximation is derived.

References

[1] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Info. Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.

[2] D. J. Love, J. Heath, R. W., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Info. Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[3] W. Santipach, Y. Sun, and M. L. Honig, "Benefits of limited feedback for wireless channels," in *Proc. Allerton Conf. on Commun., Control, and Computing*, 2003.

[4] J. C. Roh and B. D. Rao, "MIMO spatial multiplexing systems with limited feedback," in *Proc. IEEE International Conference on Communications (ICC)*, 2005.

[5] D. Love and J. Heath, R.W., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Info. Theory*, to appear.

[6] B. Mondal, R. W. H. Jr., and L. W. Hanlen, "Quantization on the Grassmann manifold: Applications to precoded MIMO wireless systems," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2005, pp. 1025–1028.

[7] W. Dai, Y. Liu, B. Rider, and V. K. N. Lau, "On the information rate of MIMO systems with finite rate channel state feedback and power on/off strategy," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2005.

[8] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and the application to MIMO systems," in *IEEE Global Telecommunications Conference (GLOBECOM)*, accepted, 2005.

[9] D. Tse and S. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimalresource allocation and throughput capacities," *IEEE Trans. Info. Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.

[10] V. Lau, L. Youjian, and T. A. Chen, "Capacity of memoryless channels and block-fading channels with designable cardinality-constrained channel state feedback," *IEEE Trans. Info. Theory*, vol. 50, no. 9, pp. 2038–2049, 2004.

[11] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc., packing in grassmannian spaces," *Exper. Math.*, vol. 5, pp. 139–159, 1996.

[12] R. J. Muirhead, *Aspects of multivariate statistical theory*. New York: John Wiley and Sons, 1982.

[13] W. Dai, Y. Liu, V. K. N. Lau, and B. Rider, "On the information rate of MIMO systems with finite rate channel state feedback using power on/off strategy," *IEEE Trans. Commun.*, submitted.

[14] A. T. James, "Normal multivariate analysis and the orthogonal group," *Ann. Math. Statist.*, vol. 25, no. 1, pp. 40 – 75, 1954.